

From People to Pixels: Visualizing Historical University Records

Tomas Vancisin tomasvancisin@gmail.com
University of St Andrews, St Andrews, KY16 9AJ, UK

Mary Orr mmo@st-andrews.ac.uk
University of St Andrews, St Andrews, KY16 9AJ, UK

Alice Crawford ac101@st-andrews.ac.uk
University of St Andrews, St Andrews, KY16 9AJ, UK

Uta Hinrichs uh3@st-andrews.ac.uk
University of St Andrews, St Andrews, KY16 9AJ, UK

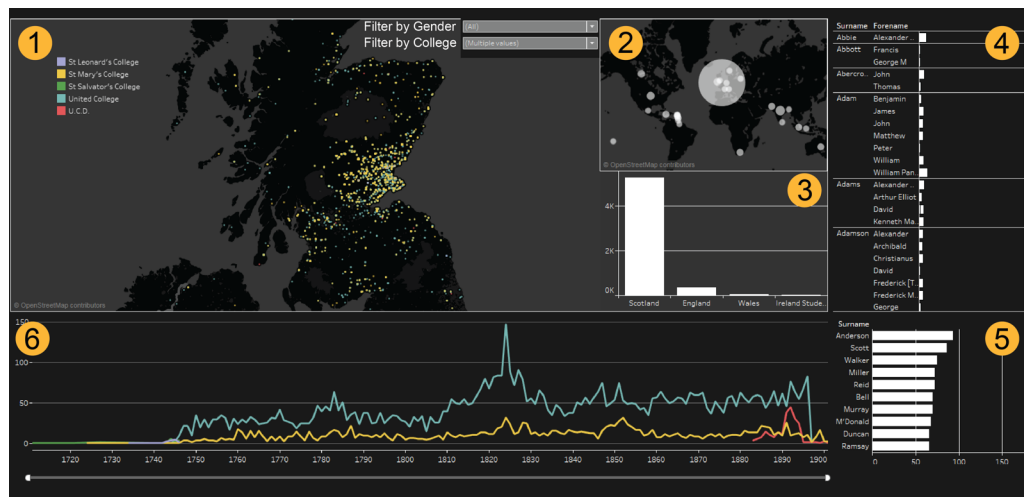


Photo: Geographical and temporal distribution of St Andrews students in different colleges 1747-1897.

Vancisin, T., Crawford, A., Orr, M., Hinrichs, U., 2018. From People to Pixels: Visualizing Historical University Records. In: Proceedings of the 5th Biennial Transdisciplinary Imaging Conference, TI2018, 18-20 April 2018, Edinburgh, UK. DOI: 10.6084/m9.figshare.6104699



From People to Pixels: Visualizing Historical University Records

Tomas Vancisin

School of Computer Science
University of St Andrews
tomasvancisin@gmail.com

Uta Hinrichs

School of Computer Science
University of St Andrews
uh3@st-andrews.ac.uk

Alice Crawford

Digital Humanities & Research
Computing
University of St Andrews
ac101@st-andrews.ac.uk

Mary Orr

School of Modern Languages
University of St Andrews
mmo@st-andrews.ac.uk

Abstract

The world's oldest universities have started digitizing their historical student/staff records [24]. Such data collections have the potential to provide valuable insights into the early educated population's social and cultural profile and inform research regarding the formation of academic networks. While textual, web-based search interfaces provide universal access to these collections for scholars and the general public, they can only provide narrow views on a record-by-record basis. This article presents and critically discusses a pilot study which uses an off-the-shelf visualization tool as a means to enable the interactive exploration of patterns within the *Biographical Register of the University of St Andrews (1747–1897)* (BRUSA). Our visualizations provide insights into the history of the University unobtainable through close reading and at the same time highlight the limitations of standard visualization tools when used in the context of diverse historical records. Drawing from ongoing advances in visualization and digital humanities (DH) research, we examine our pilot study by focusing on two main issues: (1) How to make visible the situatedness of historical (biographical) record collections? (2) How to inform the critical interpretation of cultural collections through visualization?

Author Keywords

Visualization, Digital Humanities, Historic University Records, Cultural Collections, Tableau Desktop, History

Introduction

Historical university records, a subgenre of national archives, represent lives that have shaped universities and provide insights into the development of academic research and university networks of the past. They capture information about the social and cultural background of the educated population that influenced audiences spanning from local to national, even international communities. The variety of people to whom these archives can be of great interest is therefore vast: historians, archivists, genealogists, sociologists, digital humanists as well as the general public. Whether one's purpose is to find relatives, explore the university's history, conduct research on the background of this community, or explore different approaches to record keeping, university records represent a rich source of information. Sue McKemmish describes the essential value of such collections in her article *Archives: Recordkeeping in Society*:

"The study of archives involves studying the way societies shape, hold and access information about their activities. There is no area of human activity not shaped in the most fundamental ways by the archival storage of information and no continuing form of culture or community is possible without it." [19, p.10].

Traditionally, university student and staff records were collected on paper, as part of large archives, or assembled into volumes and subsequently ordered by date and/or alphabet. Digitization of such records contributes toward preservation and makes the archives accessible to researchers as well as the interested general public. However, digitized archives, like all data collections, are only as valuable as the tools we develop to support their analysis and explorations. Web interfaces provided by the Universities

of Cambridge¹, Glasgow², Edinburgh³, and St Andrews⁴, support targeted search (by surname, degree etc.) but they cannot reveal general trends across these multifaceted collections or support their open-ended exploration.

Research in digital humanities (DH) portrays visualization—*"the use of computer supported, interactive visual representations of abstract data to amplify cognition"* [2]—as a sensible approach to characterize and analyze cultural collections by enabling interactive explorations on a large scale and from different perspectives (see Jänicke et al. for an overview of different examples [16]). Visualization can support the analysis of trends and patterns across the collection (see *"distant reading"* [20]), but it also enables transitions to small-scale subsets of the collection that might be of potential interest (see *"macroanalysis"* [17]).

We do not know of any work that focuses on the visualization of historical (biographical) university records. However, visualization approaches have been successfully applied to collections which resemble biographical archives in terms of metadata. For example, Dörk et al. have visualized temporal and professional adjacencies between philosophers and the way they have influenced each other using *EdgeMaps* [5]. *Mapping the Republic of Letters* uses visualization techniques (predominantly network and geospatial visualizations) to explore the personal and professional networks of historical scholars that emerged through travel and letter correspondence⁵. Visualization has also been discussed in the context of literary research as a means to make sense of large scale *untapped* collections where initial open-ended explorations are required to get to know the

¹<http://venn.lib.cam.ac.uk/Documents/acad/intro.html>

²<http://www.universitystory.gla.ac.uk/search/>

³<https://collections.ed.ac.uk/alumni>

⁴<https://arts.st-andrews.ac.uk/biographical-register/data/>

⁵<http://republicofletters.stanford.edu/>

Elder, William
 Student United College 1826/27 Lat. Jun.; Gk. Jun. 1827/28
 Lat. Prov.; Gk. Prov.; Math. 1; Log. 1828/29 Gk. Prov.; Lat.
 Prov.; 1829/30 Physics. Guild bursar, 1829. Part bursary,
 1828, 1829, 1830.
 Born 1804 son of William Elder, master mason, Lochee and
 Catherine Millar.
 Also at Original Secession Divinity Hall.
 Licensed by Original Secession Perth 1834.
 Minister Original Secession Church Cupar 1836 – 39. St
 Mary's, Cupar 1839 – 43. Tealing 1843 – 26.8.1890 (died).

(UC400:5.4.1828; 24.1.1829; 20.3.1830; Dundee Town
 Council minutes 2.12.1829; FES; Deaths Reg. Tealing)

Lawson, George
 Student United College 1887/88 Gk. 2; Lat. 2; Math. 1
 1888/89 Log.; Eng. Lit.; Math. 2 1889/90 N. Phil.; M. Phil.
 & Pol. Econ.; Eng. Lit.; Adv. Met.; Math. 3 1890/91 N. Phil.;
 Chem.; N. Hist.; Physiol. Lapsed bursary, 1891. Carstairs
 prize, 1889, 1890. Miller prize, 1890. Gray prize, 1891, 1892.
 Arnott prize, 1891. M.A. 30.3.1894 1st cl. hons.
 Mathematics.
 University assistant in Mathematics 1892 – 1900.
 Member of University Court 1923 - .
 Born 26.11.1870 son of George Lawson, baker, Culross and
 Janet Spittal.
 Also at Edinburgh University.
 Mathematics master George Watson's College. Rector
 Selkirk High School. Rector Waid Academy 1904 – 31
 (retired).
 Publications.
 Died in Edinburgh 15.3.1941.

(UC400:7.4.1890, 19.2.1891; UY452/21/359, 460, 500, 515,
 571; UY672; Census 1891 St Andrews; SEJ 21.3.1941;
 Deaths Reg. Edinburgh, George Square 1941/432)

Figure 1: Two BRUSA records [23,
 p.266,478]

character and facets of the collection before specifying re-
 search questions [13]. More closely related to our research,
 the l'Ecole de design Nantes Atlantique have visually ana-
 lyzed alumni career paths using networks and geospatial
 visualizations⁶. Focusing on recent data, these interactive
 visualizations show the diversity of career paths that have
 been explored by alumni. Such projects highlight the po-
 tential of visualization as a means to reveal patterns and
 unexpected insights hidden in biographical collections.

Inspired by these approaches, we have applied visualiza-
 tion as a method to explore the *Biographical Register of the
 University of St Andrews* (BRUSA), a collection of 11,894
 student and staff records spanning the period between
 1747 and 1897. While diverse in detail and depth, these
 records include student/staff name, parentage, dates and
 locations of birth and death, educational background, and
 degrees pursued at the University of St Andrews (see Fig. 1
 for two examples of such records). A textual search inter-
 face which allows users to query this collection already ex-
 exists and supports the search for student and staff records⁷
 but when it comes to large-scale explorations and analysis,
 the collection is largely untapped. Our pilot study aimed to
 investigate the potential of this collection for (1) research
 and as a method to (2) promote student, alumni, and pub-
 lic engagement with the vast history of the University of St
 Andrews (founded in 1413).

Visualizations of cultural collections are often highly cus-
 tomized to the data and context at hand which leads to
 evocative results but the process itself is rather time-consuming.
 For our pilot study, we decided to utilize a standard visual-
 ization tool—Tableau Desktop⁸—in order to rapidly design

⁶<http://justinederouet.com/alumni.html>

⁷For example, Benjamin Franklin, who was awarded an Honorary
 Doctor of Laws Degree from the University in 1759.

⁸<https://www.tableau.com/products/desktop>

a set of interactive visualizations that would allow the vi-
 sual exploration of BRUSA from multiple perspectives. The
 resulting visualizations were intended as sketches which
 would later inform the design and implementation of more
 permanent and customized visualization-based interfaces.

Our visualization sketches highlight intriguing patterns and
 confirm previous research on University history. They also
 raise interesting research questions regarding cultural and
 gender diversity at the University of St Andrews in the past,
 portraying BRUSA as a historical collection that has the
 potential to inform and guide research as well as provoke
 discussions around the history of the University among cur-
 rent students, alumni, and the general public. However, we
 also found fundamental limitations in our approach to uti-
 lize a standard visualization tool which primarily supports
 techniques that have been designed for quantitative anal-
 ysis, typically driven by empirical research [7]. Our pilot
 study shows that the resulting visualizations cannot do a
 historical collection such as BRUSA justice. This is due
 to the fact that BRUSA is diverse in focus and quality and
 situated within interpretative processes brought upon by
 archival work and digitization. The aggregated and seem-
 ingly factual quantitative views provided by our interactive
 visualization sketches gloss over questions regarding the
 transformation that collections such as BRUSA have gone
 through in terms of content, structure, representations, and
 media forms. Moreover, the individual biographies repre-
 sented by these records are hidden behind the aggregated
 birds-eye perspectives.

This paper critically outlines how visualization tools in-
 evitably influence not only the aesthetics of the visual repre-
 sentation but also the interpretation process of the underly-
 ing records. We explore possibilities for future visualization
 approaches that (1) act as commentaries on history rather

than representations of “facts”, (2) reveal and embrace the multi-layered and situated nature of historical university records, and (3) promote critical interpretation, reflection, and dialogue among academics and the general public. Our study contributes to ongoing discussions about the extent to which visualizations can be considered cultural artefacts and we also address the role of data-driven “images” in the context of cultural discourse, memory, and humanities research [7].

BRUSA

Founded in 1413, the University of St Andrews is the third oldest university in the English speaking world. A plan to publish the University’s original biographical records became real in the 19th century. The majority of the preparatory work was done by the first Keeper of Muniments, James Maitland Anderson, but the will to continue his work was inconsistent. In 1959, the first assistant to and later Keeper of Muniments at the University of St Andrews, Dr Robert Smart, started to compile a biographical register of St Andrews alumni, officers, and graduates.

Figure 2 represents an extract from one of the Graduation/Matriculation Rolls that formed the basis for Smart’s work. It includes student names, age, and information about parochial church councils. This, however, is a mere fraction of the amount of other sources Smart used over the course of 40 years during which he compiled this collection⁹. Finally published in 2004, the *Biographical Register of the University of St Andrews (1747–1897)* contains 11,894 records of former students and staff as well as external and honorary graduates. Though varying in richness

⁹Other sources include David Gregory’s lists of students, class lists, John Lee’s class records, Thomas Chalmers’ class lists, annual lists of students, separate registers, medical degree testimonials, bursary presentations, prize lists, Minutes of the Senatus Academicus, Minutes of the United College, Minutes of St Mary’s College, Library dues records.

Figure 2: Graduation/Matriculation Roll (1875–1876). [UYUY309; Image courtesy of the University of St Andrews Library]

and type of included information, records contain data such as the name, curricula, bursaries, prizes won, church affiliation, graduation dates, birth and parentage, schooling, outline careers, publications, as well as information about the sources used for information gathering (see Fig. 1 for example student records). Smart was aware of the incompleteness of this collection but his goal was to make it public and thus allow everyone to explore the history of the University based on the people who were part of it. He comments: “*The work is very incomplete and indeed never will be definitive, but I am persuaded that enough has been done to be useful and that while energy and purpose survive I should make what has been collected available to the public.*” [23]. In book form, however, public availability as well as open-ended explorations of BRUSA are highly limited.

In 2013, Digital Humanities and Research Computing teams led by Dr. Alice Crawford started to digitize BRUSA by converting the original Word manuscript into a structured, searchable XML.TEI format using manual labelling techniques (see Fig. 3). This format was indexed using an open-source enterprise search platform SOLR and subsequently converted to HTML in order to make the resulting structured data available on a web platform which allows quick searches across the entire collection (see Fig. 4). However, while making BRUSA available to interested scholars and people across the world, this platform is limited to targeted searches and does not allow open-ended explorations across the records. Such explorations are crucial from a research perspective in order to identify the potential and quality of this collection. Furthermore, from an institutional perspective, visual overviews of the collection can provide a more evocative “picture” of the history of the University.

Inspired by existing work on visualization as a method to represent large-scale biographical records presented earlier, the initial aim of our project was to design interactive visualizations that would allow a large-scale visual analysis of

```

- <tei:listPerson>
- <tei:person id="idp2bcd177ba21a41a8015ff6d8132f10df" sex="M">
- <tei:persName>
- <tei:surname>Wilson</tei:surname>
- <tei:forename>Thomas L</tei:forename>
- </tei:persName>
- <tei:education>
- Student
- <tei:orgName>United College</tei:orgName>
- <tei:date type="course" from="1880" to="1881">1880-1881</tei:date>
- Log.; Gk. 2; Lat. 2
- <tei:date type="course" from="1881" to="1882">1881-1882</tei:date>
- M. Phil. & Pol. Econ.; Eng. Lit.; Math. 2
- <tei:date type="course" from="1882" to="1883">1882-1883</tei:date>
- N. Phil.; Math. 1. Attended Free Church while a student.
- </tei:education>

```

Figure 3: XML.TEI representation a BRUSA record.

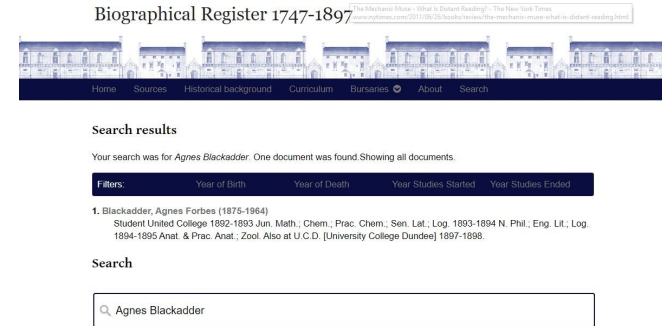


Figure 4: The BRUSA website: search results for “Agnes Blackadder”, the first female graduate of the University.

BRUSA in order to explore its research potential (relationships between academic, geographical and gender patterns etc.).

Visualizing BRUSA

First visualizations of BRUSA records can be found in Smart’s article *Some Observations on the Provinces of the Scottish Universities, 1560-1850* (1974) where he juxtaposes BRUSA records with records from Aberdeen, Edinburgh, and Glasgow [22] to allow comparisons of student numbers at Scottish universities (see Fig. 5). Beyond such quantitative views, Smart also visualized geospatial aspects inherent in student records, for example, the “zones of attraction” of individual universities in Scotland from the 17th to the 19th century (see Fig. 6).

These early visualizations motivated us to design interactive views that would allow the exploration of BRUSA in terms of gender, fields of study, and heritage (e.g., places of birth). While only representing a fraction of the information BRUSA contains, such metadata are consistently

available across all records and therefore enable rich explorations of the collection. In order to support such visualizations, we had to computationally parse the XML.TEI files for the attributes listed above and transform them into a relational database that ultimately consisted of 10 tables with attributes such as first and last name, year of birth/death, college, degree, and birth/death location.

Our visualization process was based on early paper sketches and in-team discussions about possible visual perspectives on BRUSA. We subsequently conducted a rapid visualization process using the off-the-shelf visualization tool

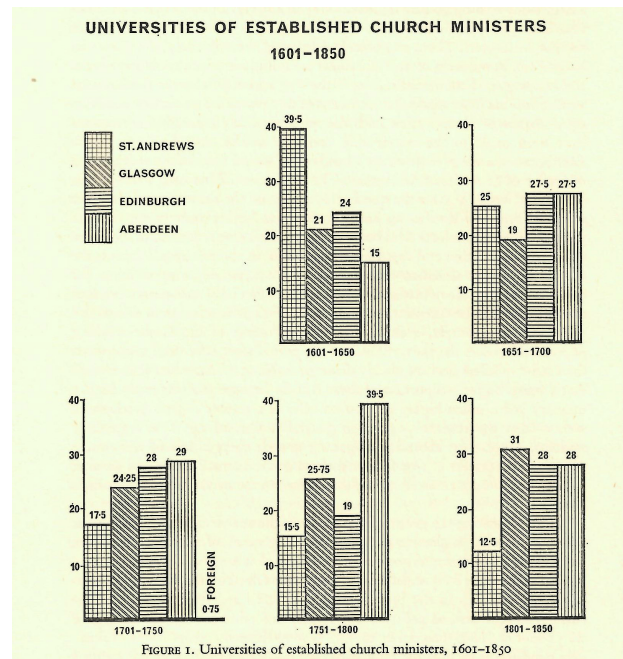


Figure 5: Universities of Established Church Ministers [22, p.95]

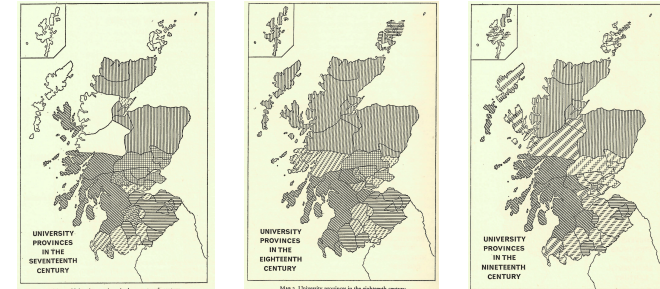


Figure 6: University Provinces in the 17th, 18th, and 19th century. [22, p.97]

Tableau Desktop. Even though this constrained our visualization design to more standard visualization techniques, it allowed us to quickly design and build visualizations of BRUSA, discuss them with scholars at the University, and engage in an initial visual exploration of this untapped collection. Below, we provide an overview of the four visualization sketches we have built as part of this process. They consist of multiple interlinked views, each focusing on a different attribute which allows their exploration in relation to each other and promote open-ended study as proposed in previous research [3, 13, 25].

St Andrews Students/Staff by Location

Visualization 1 (see Fig. 7) provides a birth location overview of students (alumni as well as those who never graduated) and staff at the University between 1747 and 1897. The visualization allows the exploration of birth locations' distribution by exact places (View 1) and by country (View 2). As most students at the time came from the United Kingdom, we provide a bar chart that displays the distribution of alumni/staff from England, Ireland (then part of the UK), Scotland and Wales (View 3). Rather than showing each

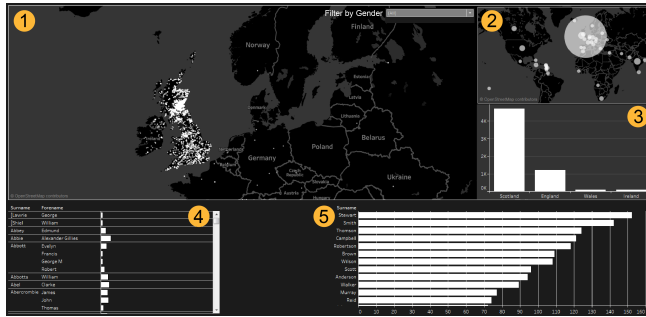


Figure 7: Vis 1: Overview of the birth locations of St Andrews students/staff.

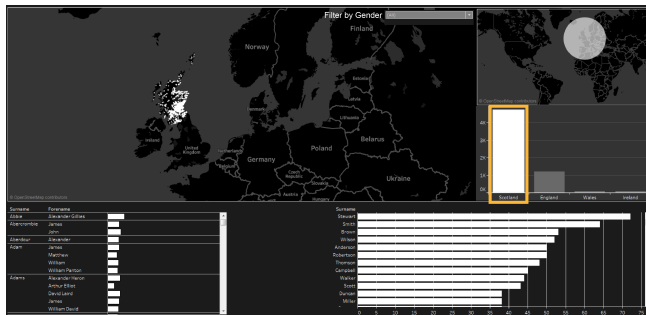


Figure 8: Selecting Scotland in View 3 adjusts the other views to only show data of students/staff born in Scotland.

record solely as a single point on the map (View 1) or in aggregated form by region (Views 2 and 3), we also wanted to present a close relation to the people behind these abstract visualizations. We therefore provided two views that focus on student and staff family names: View 4 lists all individual names across the collection as well as the amount of information available in the corresponding records (indicated by

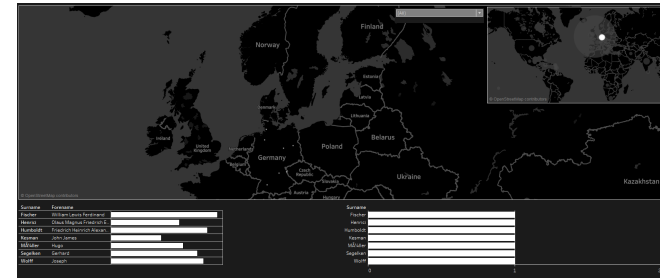


Figure 9: Filtering for German students in View 2 clears View 3.

the length of the bar next to the name). View 5 provides an aggregation of all unique last names in the collection, ordered by frequency. The five views act as a filter to one another. For example, selecting a particular region in View 3 adjusts the other views accordingly (see Fig. 8). Filtering may lead to some views being cleared completely. For example, filtering by “Germany” in View 2, will clear View 3, which only applies to regions of the UK (see Fig. 9). In addition, a drop-down filter menu is provided in View 1 which allows the filtering of records by gender (see Fig. 7).

Distribution of Internal Students by College

Visualization 2 focuses on the distribution of students (including those who never graduated) who physically attended classes in St Andrews (external students also existed) by college, in relation to their location of birth. In addition to the location and student name views as in Visualization 1, this perspective shows the temporal distribution of internal students by college and year when they started their studies (see View 6 in Fig. 10). The visualization depicts the establishment of the United College in 1747 (pale blue line) as well as the inclusion of the University College Dundee into the University of St Andrews in 1897 (red line

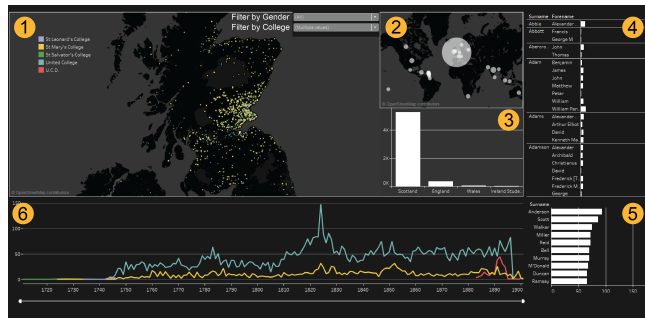


Figure 10: Vis 2: Distribution of internal students by college.

peak). Again, all views act as filters to each other. In addition, students can be filtered by gender and time period.

Distribution of Graduates by Type and Degree

Visualizations 3 & 4 (see Fig. 11 and 13) provide an overview of the degrees students acquired in St Andrews. Similarly to Visualization 1 and 2, Visualization 3 shows students' birth locations and names. An additional view presents the distribution of graduates by course type and year of grad-

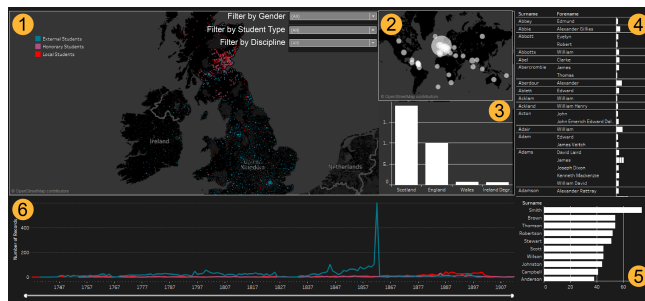


Figure 11: Vis 3: Distribution of graduates by type and degree.

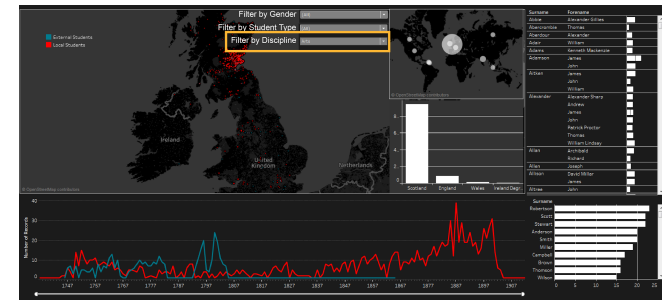


Figure 12: Filtering graduate records by Arts degree.

uation (red: internal graduates, blue: external graduates, orange: honorary graduates). Additional filters allow to focus on degree discipline (Arts, Divinity, Medicine, Law, or Science). For example, filtering by Arts degree shows that no honorary degrees were awarded in this discipline from 1747–1897 (see Fig. 12).

Visualization 4 provides more detailed perspectives on degrees awarded across the years. We categorized degrees into “sciences” and “humanities” (see Fig. 13; View 1), and further into individual disciplines, namely Arts, Divinity, Law, Medicine, and Science (View 2). View 3 shows clusters of all individual degree types (Master of Medicine [M.D.], Master of Arts [M.A.], Lady Literate in Arts [L.L.A.], etc.) with color value and rectangle size representing the number of records for each degree type. All views can be filtered by time, area of study (humanities vs. sciences), disciplines, or individual degree (see Fig. 14 & 15). We decided to juxtapose Views 1 and 2 in particular to show the distribution of degrees in complementary ways. While View 1 allows for a direct comparison of the number of degrees awarded per study area each year, generously overlooking the gaps, View 2 highlights these gaps more clearly by discipline.

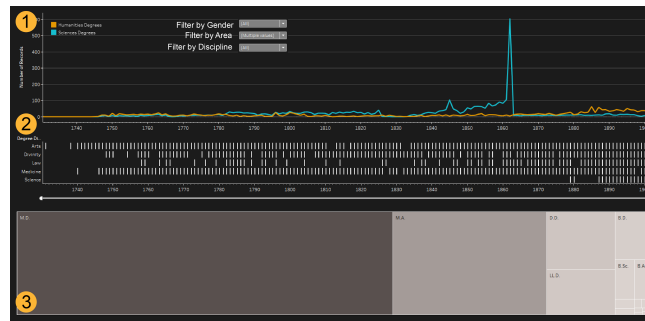


Figure 13: Distribution of degrees according to area, discipline and type.

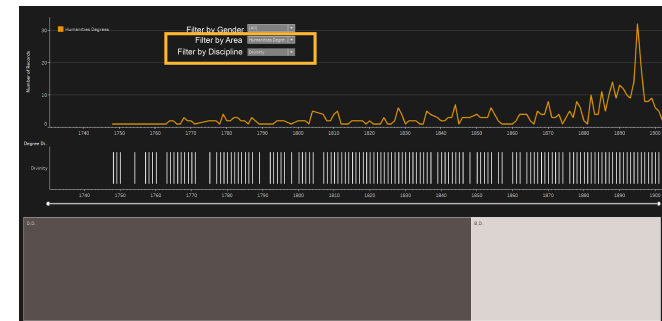


Figure 15: Filtering degrees by area and discipline: Humanities and Arts.

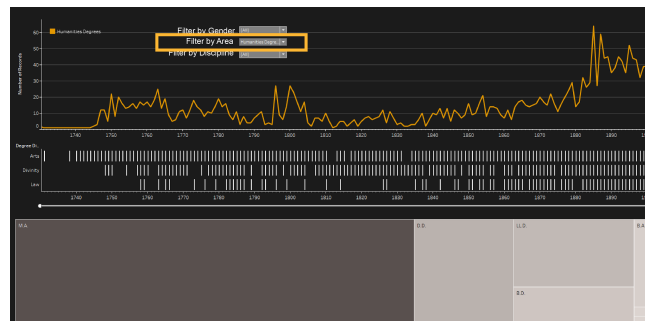


Figure 14: Filtering degrees by area of study: Humanities.

Pilot Analysis of BRUSA

The four interactive visualization sketches have provided initial insights into the history of the University and triggered additional research questions regarding geographical as well as social aspects of the records.

Geographical and Cultural Provenance

The distribution of internal students' birth locations across the world in Visualization 2 (see Fig. 10) portrays the student body of the University of St Andrews in the 18th and 19th century as international. However, the majority of internal students born outside the UK came from British colonies and only a handful were from mainland Europe. Despite the fact that only 5650 out of 7562 internal student records include information about birth locations, during the period covered in BRUSA most students who attended the University of St Andrews were linguistically or otherwise connected to the Empire. Nevertheless, it is still remarkable that some students went on such long and expensive journeys in order to obtain a university education in St Andrews.

According to the BRUSA records, the majority of students affiliated with the University (internal and external) were from all parts of the British Isles (including the Orkney islands, Hebrides, Isle of Man, Wales, England, Ireland), not only from Scotland. However, Visualizations 2 & 3 show important nuances to these distributions: students who pur-

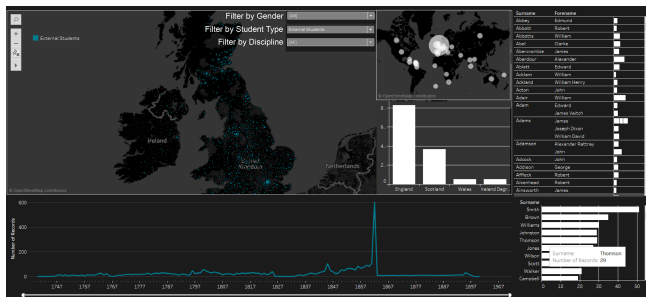


Figure 16: Distribution of external students.

sued their studies by attending classes in St Andrews were mostly Scottish (see Fig. 10). In contrast, most external students who typically gained a degree by testimonial, thesis, or examination came from England (see Fig.16)—the Uniformity Act 1662¹⁰ made it difficult for so-called “dissenters” (people who did not conform to the Church of England) to gain a degree from the old English universities such as Oxford or Cambridge [1].

Peak of Medical Degrees

Visualization 3 also revealed a compelling peak in the number of degrees awarded to external students (see Fig. 16): a sudden rise to 604 awarded degrees in 1862, followed by a massive drop to 9 degrees in the following year. Explorations with Visualization 4 show that this peak is related to medical degrees [M.D.] (see Fig. 17). Smart himself in his introduction to BRUSA explains this curious trend. The year 1863 marks the establishment of the Faculty of Medicine which required examination to earn an M.D. degree (before, the degree was awarded also by testimonial or occasionally as an honorary degree) and restricts student numbers to 10

¹⁰<http://www.british-history.ac.uk/statutes-realm/vol5/pp364-370>.

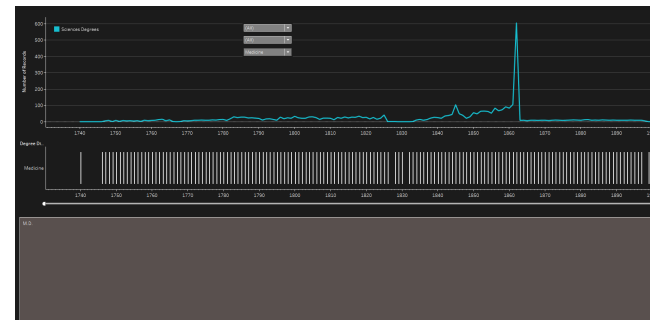


Figure 17: Peak in numbers of medical degrees.

practitioners over the age of forty per year [23, p.13]. After this change was announced, there was an apparent rush to get the M.D. before the changes took effect.

Female Students/Staff

BRUSA contains 146 female students and Visualization 3 revealed that Agnes Blackadder, known as the first female graduate at the University (1895), was neither the first female student, nor the first woman to obtain a degree in St Andrews. The visualization shows Jessie Nicholas Nelson as the first woman to graduate with the Lady Literate in Arts (L.L.A) in 1892 (see Fig. 18). This higher diploma which was of equivalent level to the M.A. was available to women from 1877. Smart writes about it in his article “*Literate Ladies—A Fifty Year Experiment*” [21] where he provides exact numbers of women who entered the program each year. Out of 699 women who started, 101 succeeded in 1892 and Nelson was one of them. The reason why she is the only one visible in the visualization for that year was divulged by Rachel Hart (current Keeper of Manuscripts at the University of St Andrews) who explains that Smart intentionally did not include any L.L.A students into BRUSA. The occurrence of 10 L.L.A.’s is incidental and can be

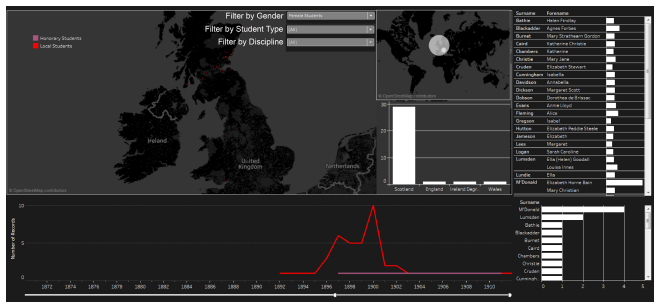


Figure 18: L.L.A. degrees included in BRUSA.

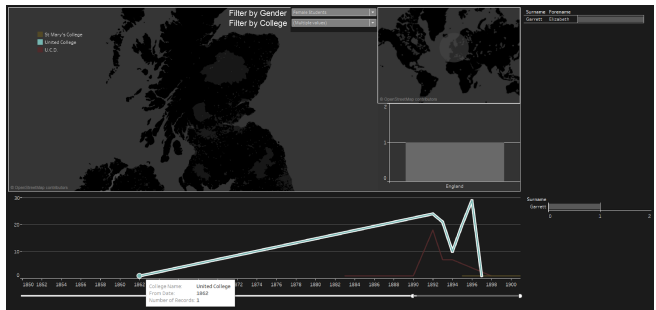


Figure 19: Elizabeth Garrett Anderson in the BRUSA records.

partially explained by the fact that some of them acquired a “proper” degree such as a M.A. later on. When taking into consideration the entire period BRUSA covers (1747–1897), 1652 women became “Literate in Arts” [21]. Including all the L.L.A.’s would not only quite dramatically change the visualizations but also present the University’s treatment of women in the 18th and 19th century in a different light. By exploring female student numbers further in Visualization 2 (see Fig. 19), another woman stands out: Eliz-

abeth Garrett Anderson was part of the United College in 1862, over ten years before the University started awarding L.L.A.’s. Additional research revealed that her matriculation was cancelled along with her applications to Oxford, Cambridge, Edinburgh Glasgow and the Royal College of Surgeons[18].

Discussion

Our pilot study provides evidence that there *is* a value to be gained from visualizing collections such as BRUSA—for scholars, the University as an institution, and the general public. Visualization can help confirm existing knowledge of the University’s history and trigger research questions that can be addressed in the future within and beyond this collection (gender diversity certainly requires an expansion of BRUSA). From an amateur’s perspective, we found that the trends highlighted by the visualizations spark interest in the history of the University and, at least for us, they triggered captivating discussions with local historians and archivists.

The visualization sketches that we developed primarily as a means to enable a rapid visual exploration of BRUSA can, of course, be improved. For example, Visualization 4 could be re-designed to allow more nuanced comparisons of obtained degrees. Furthermore, the geospatial visualizations are deceiving as we are working with a contemporary political map which is quite different from those of the 18th and 19th century. Leaving these design considerations aside, our pilot study also revealed more fundamental problems with this quantitative approach to visualizing BRUSA. Moreover, we have learned that historical (biographical) records challenge visualization as a method to facilitate interpretation and sensemaking in very specific ways.

Lawson, Andrew
Student United College 1843/44 Gk. 1; Lat. 1; Hist. Civ. 1844/45 Log; Gk. 2; Math. 1 1845/46 Lat. 2; Math. 2; Ethics 1846/47 Math. 3; Physics. Attended Free Church while a student. Lawson bursar, 1843.
Variously said to be from Aylth and from Forfar.
Also studied New College, Edinburgh – divinity student 1847/48 – 1849/50.
Licensed by Presbyterian Church of England Kyle, Northumberland.
Resident West Boldon near Sunderland from 1859.
Deleted as dead from General Council register in 1885 having not been traced since 1878.

[illegible]

cid	gender	name	id	id2	id3	id4	id5	id6	id7	id8	id9	id10	id11	id12	id13	id14	id15	id16	id17	id18	id19	id20	id21	id22	id23	id24	id25	id26	id27	id28	id29	id30	id31	id32	id33	id34	id35	id36	id37	id38	id39	id40	id41	id42	id43	id44	id45	id46	id47	id48	id49	id50	id51	id52	id53	id54	id55	id56	id57	id58	id59	id60	id61	id62	id63	id64	id65	id66	id67	id68	id69	id70	id71	id72	id73	id74	id75	id76	id77	id78	id79	id80	id81	id82	id83	id84	id85	id86	id87	id88	id89	id90	id91	id92	id93	id94	id95	id96	id97	id98	id99	id100	id101	id102	id103	id104	id105	id106	id107	id108	id109	id110	id111	id112	id113	id114	id115	id116	id117	id118	id119	id120	id121	id122	id123	id124	id125	id126	id127	id128	id129	id130	id131	id132	id133	id134	id135	id136	id137	id138	id139	id140	id141	id142	id143	id144	id145	id146	id147	id148	id149	id150	id151	id152	id153	id154	id155	id156	id157	id158	id159	id160	id161	id162	id163	id164	id165	id166	id167	id168	id169	id170	id171	id172	id173	id174	id175	id176	id177	id178	id179	id180	id181	id182	id183	id184	id185	id186	id187	id188	id189	id190	id191	id192	id193	id194	id195	id196	id197	id198	id199	id200	id201	id202	id203	id204	id205	id206	id207	id208	id209	id210	id211	id212	id213	id214	id215	id216	id217	id218	id219	id220	id221	id222	id223	id224	id225	id226	id227	id228	id229	id230	id231	id232	id233	id234	id235	id236	id237	id238	id239	id240	id241	id242	id243	id244	id245	id246	id247	id248	id249	id250	id251	id252	id253	id254	id255	id256	id257	id258	id259	id260	id261	id262	id263	id264	id265	id266	id267	id268	id269	id270	id271	id272	id273	id274	id275	id276	id277	id278	id279	id280	id281	id282	id283	id284	id285	id286	id287	id288	id289	id290	id291	id292	id293	id294	id295	id296	id297	id298	id299	id300	id301	id302	id303	id304	id305	id306	id307	id308	id309	id310	id311	id312	id313	id314	id315	id316	id317	id318	id319	id320	id321	id322	id323	id324	id325	id326	id327	id328	id329	id330	id331	id332	id333	id334	id335	id336	id337	id338	id339	id340	id341	id342	id343	id344	id345	id346	id347	id348	id349	id350	id351	id352	id353	id354	id355	id356	id357	id358	id359	id360	id361	id362	id363	id364	id365	id366	id367	id368	id369	id370	id371	id372	id373	id374	id375	id376	id377	id378	id379	id380	id381	id382	id383	id384	id385	id386	id387	id388	id389	id390	id391	id392	id393	id394	id395	id396	id397	id398	id399	id400	id401	id402	id403	id404	id405	id406	id407	id408	id409	id410	id411	id412	id413	id414	id415	id416	id417	id418	id419	id420	id421	id422	id423	id424	id425	id426	id427	id428	id429	id430	id431	id432	id433	id434	id435	id436	id437	id438	id439	id440	id441	id442	id443	id444	id445	id446	id447	id448	id449	id450	id451	id452	id453	id454	id455	id456	id457	id458	id459	id460	id461	id462	id463	id464																																																																																																																																																																																			
cid006...	M	Victor Ale...	...	Bruce	...	43	1849																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														

Figure 20: Transformations of BRUSA records.

Making Archival Processes Visible

D'Ignazio and Klein in their article *Feminist Data Visualization* (2016) emphasize that making the “labour” visible gives credit to the people who collected, digitized, and maintained the data while highlighting “*fair attribution and credit for the resulting artifact*” [15]. For collections such as BRUSA, this is especially crucial because they undergo multiple stages of interpretation manifested not only in their content but also in their structure, (visual) representation, and form of presentation (as illustrated in Fig. 20). It is important to remember that all record keeping and archiving processes are affected by the subjectivity of their proponents. As Hillary Jenkinson observes in her article *Reflections of the Archivist* [11, p.20], “operative bias” must always be kept in mind when examining the archival process:

"Perhaps one of the most difficult parts of this task is the guarding against [the archivist's] own interest in anything except their safety. The Archivist must not turn Student, or may at most do so only as an occasional treat, and with the strictest precautions against his own possible malfeasance; for every Student has an axe to grind, a theory to establish, a statement to prove: and that form of interest is incompatible with dispassionate conduct in sorting, in arrangement, in presentation-in all those processes in which the tiniest modification may have the most far-reaching results."

Smart's meticulous process which involved careful, yet selective archival and curatorial work (e.g., evident in his decision to exclude L.L.A. graduates) (re-)structured the biographical records and introduced consistent terms of description. By setting the records in a modern text editor, he transformed the visual representation and media form in which the records are presented. But decisions about structure and context of the records did not end with Smart's publication of BRUSA. The TEI tagging which involved

the manual identification of key content elements in each record further transformed the records' structure which allowed new representation and media forms (e.g., a searchable web interface) to emerge. Last but not least, our own approach of computationally parsing the XML-TEI files and storing the resulting metadata in a relational database introduced additional transformations in terms of structure and (visual) representation as well as content changes (mostly characterized by omissions of record aspects that we have yet to computationally parse). Again, all these transformations have had a fundamental effect on BRUSA's content, the ways in which it is presented, and most importantly how the collection is interpreted. While the ancient manuscripts immediately show the diversity of primary sources that feed this collection and hint at the interpretation and curation necessary to compile BRUSA, our visualizations do not. This is problematic because all changes that BRUSA records underwent can be considered as (historical) layers of interpretation driven by different intentions and constraints. For example, while Smart's and Anderson's decisions were driven by their perspective as historians and archivists, our pilot study was fundamentally shaped by technological and design considerations. This raises crucial questions: How to make such layers of transformation and interpretation visible? How can we enable fluid transitions between abstract visualizations we have built and the original data material that "drives" them?

Previous research has highlighted that the loss of material qualities, as part of digitization, is a problem for historical or literary collections [8] and our argument goes in a similar direction. Juxtaposing different textual and visual forms of representations in a collection such as BRUSA could be one “avenue” to explore (as illustrated by [13, 26]). Another possibility would be the visual and interactive integration of different forms of media in which these historical collections

have been represented (including pictures of manuscripts that highlight their original quality) into visualizations of historical collections (as illustrated by [4, 9, 27]).

Historical Data Uncertainty

BRUSA is a highly diverse collection of biographical records and its incompleteness is emphasized by Smart himself [23]. Some of the records contain detailed information about individuals affiliated with the University of St Andrews while others are more sparse. For example, only 6339 records contain birth locations of staff/students, hence 47% of BRUSA records are not visible in the geospatial views shown in Visualizations 1, 2, and 3. In fact, the quantitative overviews we have designed provide only a seemingly objective representation of the collection which, without a background of the record collection and how it has been processed, can easily be interpreted as an overview of “all staff and students at the time”. While the visualizations can act as entry points into research questions, on their own they cannot provide definitive answers and require additional research.

Visualizations come with their own rhetorical qualities [14]. Johanna Drucker argues that standard visualization techniques (line graphs, bar charts), traditionally designed for and used in the empirical sciences, are now being used in the context of humanities research which runs the risk of presenting information in a familiar and seemingly definitive way while hiding the assumptions, underlying interpretations, and decisions that influenced the data collection, categorization, and finally, the visualization itself [7]. Drucker highlights that “*the basic categories of supposedly quantitative information, the fundamental parameters of chart production, are already interpreted expressions. But they do not present themselves as categories of interpretation, driven with ambiguity and uncertainty, because of the representational force of the visualization as a “picture”*

of “data”.” [7]. Our visualization sketches are an illustration of this problem. They highlight that while standard visualization tools allow the rapid exploration of multi-faceted cultural collections, it is absolutely crucial to critically discuss them not as a means to an end but as the product of an interpretative, speculative process (see the discussion of “visualization sandcastles” by Hinrichs and Forlini [12]) which can inform the design of customized visualizations capable of revealing more from the records’ rich historical and interpretative layers as well inherent uncertainties. The area of visualizing uncertainty is vast but we highlight work by Wood et al. on “sketchy rendering for information visualization” as an interesting approach for further design explorations in this area [28].

Showing the People Behind the Data

Our visualization sketches provide highly aggregated views of thousands of University students and staff members, merging the individual into a homogeneous mass. While useful for the exploration of patterns in the data, this birds-eye perspective may alienate the viewer from what this collection also represents: individuals, each with their own unique lived experience that is connected to the University of St Andrews. We tried to keep the individuals behind the aggregated views present in our visualization sketches through a scrollable list of names which could also provide access to their individual records. However, this view, quantified through a bar chart, still fails to account for the crucial distinctiveness of each person represented in the records. This is problematic when trying to use collections such as BRUSA as a means to make history more graspable or help imagine the life of an individual at the University in 18th/19th century Scotland.

How can we (1) communicate that the trends shown in our visualizations are formed by distinctive characters whose

paths have crossed in St Andrews and (2) how can we enable explorations of BRUSA through a birds-eye view as well as the lens of individual student/staff biographies? Addressing these questions from a design perspective, Dörk et al. have introduced the idea of “monadic interaction” which “*brings the contradictory representations of individual elements and entire collections closer together and suggests continuous movements between partially overlapping vantage points.*” [6]. Applying the idea of monadic interaction to collections such as BRUSA would mean to give each biography its own presence in the visualization (e.g., in the form of an individual element such as a circle, where size represents the amount of information in the record and a slightly irregular shape represents distinctness from other biographies). Interacting with this element could then highlight connections to other biographies in the form of other visual elements. Additionally, the famous, better-known alumni could be brought up to initiate and guide the navigation of the records and clustering could be applied as a middle ground between aggregated visualizations and the representation of every individual biographical record.

Conclusion

“Of course, one purpose of history might just be to bridge the gap between past and present, restoring this ruptured continuity. But how can currents of collective thought whose impetus lies in the past be re-created, when we can grasp only the present?” Maurice Halbwachs, 1992 [10]

Despite the fact that it is impossible to build a genuine “bridge” to the past, the combination of visualization and humanities research techniques provides possibilities to bring the shores of past and presence closer together, and our study aims to do exactly that. It represents a starting point into exploring ancient archives through perspectives that would be impossible to gain using traditional archival and close read-

ing methods. At the same time, the visualization sketches we developed have revealed that in order to grasp the full potential of historical collections that represent thousands of individual biographies, critical interpretation of visualization artifact and thinking beyond traditional visualization tools and toward customized visualizations are a “must”. To elicit critical interpretation of such diverse records, it is crucial to take into consideration (1) the transformation processes these collections underwent, (2) acknowledge the uncertainty and situated interpretations inherent in these records and (3) give individual records (in our case human biographies) a presence in all their distinctiveness, instead of prioritizing aggregated overviews that easily gloss over what makes historical university records so unique.

Acknowledgements

First and foremost, we would like to thank Robert Smart whose mind-bending dedication to compile BRUSA keeps fuelling this project with inspiration. We also want to express our gratitude to Rachel Hart and the Special Collection Department as well as the University Library and its Digital Humanities and Research Computing teams for their ever-present support.

References

- [1] M. G. Brock. 1997. *Nineteenth-century Oxford*. Clarendon Press.
- [2] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman. 1999. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers.
- [3] M. Dörk, S. Carpendale, C. Collins, and C. Williamson. 2008. VisGets: Coordinated Visualizations for Web-based Information Exploration and Discovery. *IEEE TVCG* 14, 6 (2008), 1205–1212.
- [4] Marian Dörk, Sheelagh Carpendale, and Carey Williamson. 2012. Fluid Views: A Zoomable Search

- Environment. In *Proc. of AVI*. 233–240.
- [5] Marian Dörk, Sheelagh Carpendale, and Carey Williamson. 2012. Visualizing explicit and implicit relations of complex information spaces. In *Information Visualization*, Vol. 11. Sage, 5–21.
- [6] Marian Dörk, Rob Comber, and Martyn Dade-Robertson. 2014. Monadic Exploration: Seeing the Whole Through Its Parts. In *Proc. of CHI*. 1535–1544.
- [7] J. Drucker. 2011. Humanities Approaches to Graphical Display. *Digital Humanities Quarterly (DHQ)* 5, 1 (2011).
- [8] Stefania Forlini and Uta Hinrichs. 2017. Synesthetic Visualization: Balancing Sensate Experience and Sense Making in Digitized Print Collections. In *Proc. of DPASSH*.
- [9] Katrin Glinka, Christopher Pietsch, and Marian Dörk. 2017. Past Visions and Reconciling Views: Visualizing Time, Texture and Themes in Cultural Collections. 11 (01 2017).
- [10] Maurice Halbwachs. 1992. *On collective memory*. University of Chicago Press, Chicago.
- [11] Hillary Jenkinson. 1984. Reflections of an Archivist. In *A Modern Archives Reader*, Maygene F. Daniels and Timothy Walch (Ed.). National Archives and Records Service, Washington D.C., 15–21.
- [12] Uta Hinrichs and Stefania Forlini. 2017. In Defense of Sandcastles: Research Thinking through Visualization in DH. In *Proceedings of Digital Humanities*.
- [13] Uta Hinrichs, Stefania Forlini, and Bridget Moynihan. 2016. Speculative Practices: Utilizing InfoVis to Explore Untapped Literary Collections. *IEEE TVCG* 22, 1 (2016), 429–438.
- [14] Jessica Hullman and Nicolas Diakopoulos. 2011. Visualization Rhetoric: Framing Effects in Narrative Visualization. *TVCG* 17, 12 (2011), 2231–2240.
- [15] Catherine D Ignazio and Lauren F Klein. 2016. Feminist Data Visualization. In *Proceedings of the Workshop on Visualization for the Digital Humanities (Vis4DH)*, held in conjunction with IEEE VIS.
- [16] S. Jänicke, G. Franzini, M.F. Cheema, and G. Scheuermann. 2016. Visual Text Analysis in Digital Humanities. *Computer Graphics Forum* 36, 6 (2016), 226–250.
- [17] Matthew Jockers. 2013. *Macroanalysis : digital methods and literary history*. University of Illinois Press, Urbana.
- [18] Jo Manton. 1965. *Elizabeth Garrett Anderson*. Methuen, London.
- [19] Sue McKemmish. 2005. *Archives : recordkeeping in society*. Centre for Information Studies, Charles Sturt University, Wagga Wagga, N.S.W.
- [20] Franco Moretti. 2013. *Distant Reading*. Verso.
- [21] Robert N. Smart. 1968. Literate Ladies - A Fifty Year Experiment. *St Andrews University Alumnus Chronicle* 59 (1968), 21–31.
- [22] Robert N. Smart. 1974. Some observations on the provinces of the Scottish Universities, 1560-1850. *The Scottish Tradition* (1974), 91–106.
- [23] Robert N. Smart. 2004. *Biographical Register of the University of St Andrews, 1747-1897*. St Andrews University Library.
- [24] Bernhard Spörlein, Akademie (Bamberg), Gesellschaft für Fränkische Geschichte., and Universität Bamberg (1773-1803). 2014. *Die Matrikel der Akademie und Universität Bamberg 1648 - 1803. 2, Personen- und Ortsregister*. Ges. für Fränkische Geschichte. <http://genwiki.genealogy.net/Universitätsmatrikel>
- [25] Alice Thudt, Uta Hinrichs, and Sheelagh Carpendale. 2012. The Bohemian Bookshelf: Supporting Serendipitous Book Discoveries through Information Visualization. In *In Proc. of CHI*. 1461–1470.

- [26] Jim Clifford Andrew Watson Aaron Quigley Ewan Klein Colin M. Coates Uta Hinrichs, Beatrice Alex. 2015. Trading Consequences: A Case Study of Combining Text Mining & Visualisation to Facilitate Document Exploration. *Digital Scholarship in the Humanities (DSH)* 30, 1 (2015), i50–i75.
- [27] M. Whitelaw. 2015. Generous Interfaces for Digital Cultural Collections. *Digital Humanities Quarterly (DHQ)* 9, 1 (2015).
- [28] Jo Wood, Petra Isenberg, Tobias Isenberg, Jason Dykes, Nadia Boukhelifa, and Aidan Slingsby. 2012. Sketchy rendering for information visualization. *IEEE TVCG* 18, 12 (2012), 2749–2758.